

WHAT IS CLAIMED IS:

1. A method of identifying a characteristic of interest represented by a textual input, comprising:
building a graph corresponding to the textual input;
scoring sub-graph components of the graph;
identifying graph fragments of interest based on the scores; and
performing text manipulation based on the identified graph fragments.
2. The method of claim 1 wherein building the graph includes building the graph with nodes linked by links.
3. The method of claim 2 wherein the nodes correspond to words in the textual input or concepts represented by the textual input.
4. The method of claim 3 wherein building the graph further comprises generating the links as directed semantic relation names.
5. The method of claim 4 wherein building the graph further comprises generating a set of abstract analyses for the textual input.
6. The method of claim 5 wherein generating a set of abstract analyses comprises:

generating a set of directed acyclic graphs
based on the textual input; and
connecting the set of directed acyclic graphs to
one another.

7. The method of claim 2 wherein building the graph
comprises:

generating a syntactic parse for text portions
in the textual input;
generating a dependency structure from the
syntactic parse; and
generating the graph from the syntactic parse.

8. The method of claim 2 wherein building the graph
comprises:

identifying the nodes as adjacent or collocated
words; and
identifying the links between the nodes.

9. The method of claim 8 wherein identifying the
links comprises:

assigning directionality of the links
arbitrarily.

10. The method of claim 8 wherein identifying the
links comprises identifying the links and assigning
directionality of the links based on a given part-of-
speech associated with the nodes, using a heuristic.

11. The method of claim 8 wherein identifying the links comprises identifying the links and assigning directionality of the links based on a given part-of-speech associated with the nodes, using a machine learned method.

12. The method of claim 2 wherein scoring sub-graph components of the graph comprises:

assigning a score to each node in the graph.

13. The method of claim 12 wherein a pair of nodes and a link between the pair of nodes comprises a tuple and wherein scoring sub-graph components of the graph comprises:

assigning a score to each tuple in the graph.

14. The method of claim 13 wherein assigning a score to each tuple in the graph comprises:

assigning a score to each tuple in the graph based on a score of an initial node in the tuple, scores for nodes linking to a target node in the tuple, and a frequency of the tuple in the textual input.

15. The method of claim 13 wherein identifying graph fragments of interest comprises:

matching sub-graph components of the graph to nodes and tuples having a sufficient scores.

16. The method of claim 15 wherein identifying graph fragments of interest comprises:

identifying nodes, having a sufficient score,
that are linked to the matched sub-graph
components.

17. The method of claim 16 wherein identifying graph fragments comprises:

identifying a node outside a matched sub-graph
component that has a predetermined relation
to a node in the matched sub-graph
component.

18. The method of claim 17 wherein identifying graph fragments comprises:

identifying certain relations, given a
predetermined specific node type.

19. The method of claim 18 wherein all the matched sub-graph components and identified nodes and relations comprise the graph fragment.

20. The method of claim 19 wherein performing text manipulation comprises:

extracting the set of sub-graph components
identified for a given portion of the
textual input as a graph fragment.

21. The method of claim 20 wherein building a graph comprises:

generating a separate graph for each sentence in the textual input; and
connecting the separate graphs together to form an overall graph.

22. The method of claim 21 wherein extracting comprises:

extracting sub-graph portions, that have a sufficient score, from the overall graph.

23. The method of claim 21 wherein high scoring sub-graph portions of the overall graph comprise sub-graph portions of the overall graph that have a score that meets a threshold score value, and wherein extracting sub-graph portions comprises:

extracting portions of the separate graphs that spawned the high scoring sub-graph portions of the overall graph.

24. The method of claim 1 wherein performing text manipulation comprises one of summarization, information retrieval, question answering, document clustering, and indexing.

25. The method of claim 1 wherein performing text manipulation comprises: generating a textual output based on the extracted graph fragments.

26. The method of claim 1 and further comprising:

ordering the graph fragments based on scores corresponding to the graph fragments.

27. The method of claim 26 wherein ordering further comprises:

ordering the graph fragments based on factors in addition to the scores.

28. The method of claim 27 wherein the factors comprise one of placement of nodes and the order in which two nodes related through part of speech will occur, an event timeline determined from the textual input, and a topic determined for the textual input.

29. The method of claim 1 wherein the characteristic of interest comprises one of words, text fragments, concepts, events, entities, and topics.